Developing Language Test: Combining Sentences

Keiko Sato

INTRODUCTION

Frequently language teachers are and have been testing the student achievement level for several purposes; to evaluate what students have learned from a certain course, to get the information of what they know and what they do not, to measure the development of students' language ability in a certain period, to estimate the effectiveness of a certain kind of teaching in a given time, and so on. Most of those teachers make test items for themselves without thinking about reliability, validity, or objectivity. They believe their experience and subjective intution. Consequently, the fact is that no one knows whether they measure what they want to measure properly.

Researchers of language acquisition are in more grave position. They cannot go on with their research without relying on the reliability and validity of tests they intend to use. They always worry whether the test will be a good measuring tool when they are going to use it. Some of them, of course, may test the reliability and validity of the test before using it, and try to remove factors affecting the reliability and validity and revise the test items. It is true, however, that most of the others are using tests leaving their disbelief in them. Both language teachers and language learning researchers must know, first of all, the distinction between norm-referenced tests (NRTs) and criterion-referenced tests (CRTs). NRTs have dominantly been used for evaluating the testee's proficiency of a language in terms of the performance of the others, that is, the normative group. The representative is TOEFL widely used in the world. Recently the problems with NRTs such as 'teaching mismatches, lack of instructional sensitivity' (Brown and Hudson 1992c) and many others led to the development of CRTs, which provide information of the degree of the testee's acquired ability in terms of his or her performance without reference to the others'. The tester needs to decide which type of test is suitable to the purpose for which the test is undertaken.

This paper is attempting to develop CRTs according to T. Hudson's format. The writer has been interested in helping students develop writing English in length, and has been investigating what they have learned and what they have not yet by having them write combined sentences. It seems necessary at this point that those types of test and the way of generate test items should be considered in the light of test specifications and item specifications.

DEVELOPING LANGUAGE TEST

I.　Description of the Context of the Test

A.　Skills to Be Tested

This test is designed to evaluate the students' ability of expanding sentence by combining words, phrases and clauses. It can be used for a pretest or post test,and also used as an achievement test. Grammatical devices to be tested include "adjectives," "prepositional phrases," "verbals," and "embedded clauses".

B. Audience

The examinees expected are the freshmen (actually women) in the composition course of a Japanese college. They all have six years of experience of studying Engishin the formal school education. However, they are not expected to have written anypassage in English through their own ideas. It is taken for granted that teaching composition is taught in the manner of translation.

C. Course in Which the Test is Given

The test is given to the students the administrator herself teaches. She is in charge of two classes, the sizes of which are 52 and 54. The objectives of the composition course she made are that the students will write English spontaneously without hesitation and with no care of minor mistakes and that their writings will come close to the length and complexity of sentences written by native speakers of the same age.

II. Assessment Methods

A. Description

The tests are given before, during, and at the end of, the course to measure the level of the students' performance, and evaluate the effect of teaching writing ways using specific grammatical devices. A test of this kind may be called "criterion-referenced test" (CRT) in that this test "is primarily designed to describe the performances of students in terms of the amount that has been learned of a specific domain of knowledge or set of" skills. (James Dean Brown and Thom Hudson c 1992)

B. Rationale for the New Test

A test of this kind has the effect of both evaluation and consciousness raising. ESL/EFL learners do not notice that there is a proper length and complexity of sentence syntactically and stylistically. In the reading class setting, they turn their eye to the sentence structure, whereas in the writing course they pay attention to grammar and content they intend to express.

According to Kellog W. Hunt (1965), American adults who like to write make their sentences 24.7 words in average and students in the twelve grade write with a mean length of 16.9 words, while according to Sato's (1991) investigation, Japanese students of the same grade write sentences of 12.76 words. Judging from the lack of consciousness raising of sentence expansion devices, it is expected that there is some possibility of developing students' writing in this dimension. By teaching each grammatical device discretely first and having their gained ability integrated in their writings, their resulting product will be one similar to native speakers' products. In this respect it is necessary to make the CRT model for combining sentences in EFL classes in Japan.

III. Test Specification

A. Overall Test Descriptor

This is a paper and pencil test to determine mastery of a specific technique for writing English. Even if the examinees of English composition course demonstrate mastery of the specific devices, they will not be placed in the upper class individually. The advantage is just the fact that their writings are actually improving in that specific aspect and the satisfaction that they are writing better sentences. The test covers grammar and writing components which assess logical, stylistic and linguistic accuracy abilities. Examinees who demonstrate mastery minimally meet the following profiles:

1. *Logical*:   The examinee demonstrates mastery of: 1) using grammatical devices logically (order of clauses, use of connectors); and 2) coherent building of discourses (use of transitional devices)

2. *Stylistical*:   The examinee demonstrates mastery of appropriate length of sentences involved in appropriate complexity.

3. *Accurate*:   The examinee demonstrates mastery of advanced structures of English grammar in context. Structures include embedded forms, modifications, subordinations and coordinations.

B. General Level Descriptors

This test does not determine examinees' levels of mastery, but gives information of what devices an individual examinee masters or what devices are difficult for an examinee to employ. It also tells both an examinee and his/her instructor that the examiner reaches to what extent he/she acquires a particular device for producing an appropriate length of sentence.

C. Specific Test Descriptors

Grammar: A person who masters this area is required to demonstrate performances of advanced grammatical structures in two to four sentence contexts. Structures tested include:

embedded forms:
Subbordination
Embedded Clauses
Noun Clauses
Adjective Clauses
Adverb Clauses
Embedded Phrases
Infinitives
Participles
Gerunds
Prepositional Phrases

Coordination

Coordinated Words

Coordinated Phrases

Coordinated Clauses

Writing: A person who masters this area is required to demonstrate appropriate length of sentences in coherent meaningful discourses. Tasks in this area:

Controlled Writing

Making a story describing a series of pictures, continuing a paragraph after a topic sentence is given.

IV.　Item Specification

Combining Sentences

A.　*General description:*

When presented with two or more sentences, the student will combine them into one without missing any information by using some grammatical devices.

B.　*Sample Item:*

Directions: The following questions (1-10) have several sentences, followed by some connectives. Combine those sentences into one, using the connectives given.

Example:　　I slept in the sun all afternoon.

I got a terrible sunburn.　　　　　(so)

C.　*Prompt Attributes:*

1. Each sentence will contain only one or very few pieces of information in a simple structure.

2. Each question will have two or more possible sentences to join together.

3. The whole test will cover as many different grammatical devices as possible.

D.　*Response Attributes:*

1. The student will combine some sentences given into one using the word(s) given in the space provided on the answer sheet.

2. The correct response will be a sentence which:

a.　includes the information of original sentences.

2 point: Mastery: sentence expresses all the information the original sentences convey.

1 point: Semi-Mastery: sentence expresses most of the original information, but lacks one or two pieces of minor information.

0 point: No Mastery: sentence lacks more than two pieces of information.

b.   uses the grammatical devices given.

2 point: Mastery: the connectives or grammatical hints given are used correctly.

1 point: Semi Mastery: the connectives given are used correctly except one connective.

0 point: No Mastery: two or more connectives are used incorrectly.

E.   *Specification Supplement*

Connectives to be employed are as follows:
   simple subordinators:
      after, (al)though, as, because, before, if, of, once, since, that, until, when, where, while, etc.
   compound subordinators:
      in that, so that, except that, in order that, now (that), provided (that), supposing (that), as far as, as long as, as if, as though, in case, etc.
   correlative subordinators:
      as...so, more...than, so...that, such...that, no sooner...than, etc.
   simple coordinators:
      and, but, or, so, for, etc.
   correlative coordinators:
      both...and, either...or, neither...nor, not...but, not only...but also, etc.
   relatives:
      that, who, whose, whom, which, what, but, where, when, why, etc.
   non-finite postmodifiers:
      -ing participles, -ed participles, infinitives, prepositional phrases, appositives, etc.

Controlled Writing

A.   *General Description:*

   When presented with a series of pictures or the first topic sentence, the student will make his/her own story and write a paragraph in English.

B.   *Sample Item:*

Directions : Look at the pictures carefully. This story is about a young man named Mr. Sato. Write your story about him, referring to the pictures. Begin with the sentence given below the picture.

Complete each story in 5 to 10 sentences.

Example:
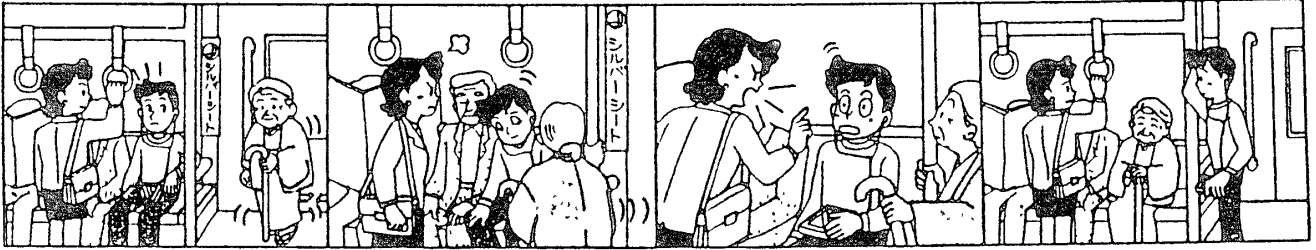
One morning Mr. Sato was riding on the train.

C. *Prompt Attributes:*

1. Pictures will be reprinted from Eiken semi-1st degree tests already given and published by the Association of Japan English Education.

2. Topics will be accessible fo students so that their ideas will soon come to them.

3. Starting sentences will be provided for the student to start writing with ease.

D. *Response Attributes:*

1. The student will write five to ten sentences within five minutes for each question.

2. The correct answer will be:

   a. for grammar: writing with complex structures. It is better to use as many embedding devices as possible.

   b. for length: writing with lengthy sentences. It is preferable that sentences consist of more than twenty words.

V. Items

Sentence-Combining Test

Directions: Each of the following questions (1-10) presents several sentences, followed by some connectives or grammatical hints in the parentheses. Combine those sentences into one, using the conectives or hints given. You may delete an element or change the form of an element if necessary.

Example A: I slept in the sun all afternoon.

I got a terrible sunburn.   so

The answer is "I slept in the sun all afternoon, so I got a terrible sunburn." So write it in the space provided on the answer sheet. You will have 20 minutes for 10 questions.

Example B: I have a book.

I bought it yesterday. (relative pronoun)

The answer is "I have the book which (or that) I bought yesterday." So write it in the space provided on the answer sheet. You will have 15 minutes for 10 questions.

1. Hiroko married a man.

The man is poor.   (preposed adjective)

2. I rode my bicycle.

I went over the bridge.   (postposed

I went to the museum.   prepositional phrase)

3. Many students work hard as part-time workers.

They want to get extra money.   (infinitive)

4. He arrived at the station.

He found that the train had already left.   (participle)

5. His suit was ruined by the rain.

It had to be thrown away.   (because)

6. She took out a hundred yen.

She had been saving it.   (relative pronoun)

7. Sumiko walked into the game center.   (when)

Sumiko is a disco crazed college student.   (appositive)

Sumiko is holding her boyfriend's arm.   (participle)

8. The Olympic Games start.   (before)

An athlete appears.

He holds a torch.   (participle)

It has been carried from Mount Olympus in Greece. (relative pronoun)

9. The athlete enters the stadium.   (when)

He is holding the torch.   (relative pronoun)

He runs to the huge bowl.   (and)

The sacred flame will burn there.   (relative pronoun)

10. The teacher is not beautiful.   (though)

　　 The teacher is not great.   (neither...nor)

　　 The teacher has blue eyes.   (prepositional phrase)

　　 Most students like the teacher.

Controlled Writing

Directions: Look at the pictures carefully. The story of Question 1 is about the Yamadas, and the story of Question 2 is about Mr. Tanaka, an office worker. Write your story for each question, referring to the pictures. Begin with the sentence given to each story below the pictures. You will have 10 minutes for two questions.
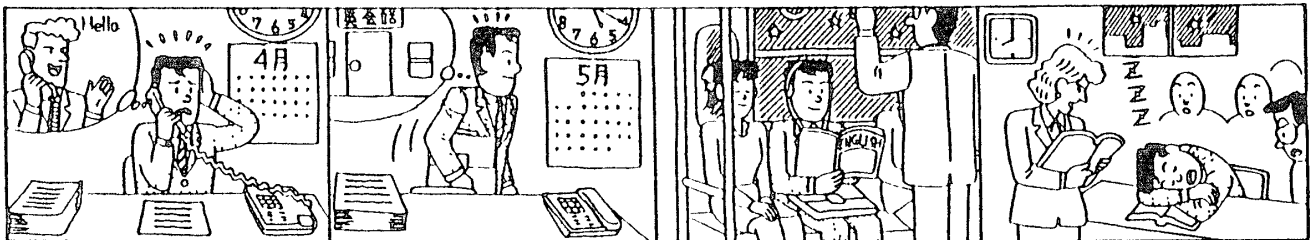
Question 1

One day the Yamadas went to the zoo._____

_____

_____

_____

_____

Question 2

Mr. Tanaka joined the ABC Trading Company in April._____

_____

_____

_____

_____

_____

VI. The Cognitive Processes Involved in Each Item

A. Task Dependence

Both of our tests are task-dependent. The testee must produce his/her own sentences. Receptive knowledge is not helpful to this kind of test. Producing new sentences requires analytic and synthetic language abilities. There is little room for inference to come in.

With the combining test, the pure productive ability will not be measured, but syntactically manipulating ability can be presented in the product, which is congruent to the test specification. The objective of this test is indeed to realize the degree to which the testees can combine some pieces of information into one structure.

With the picture description test, there are several abilities involved, such as inference from the pictures, imagination for story making, construction of utterly new sentences. It requires integrated language skills.

B. Production Skills

This test focuses on a specific manipulation of syntactic structure; producing complex structure with or without finite verb. This aim will be productive in a sense, though the combining test requires less productivity than the picture description test. EFL learners should take this learning process.

C. Reasoning, Inferencing Ability

This test requires reasoning ability especially when two or more pieces of information are connected with subordinators. The relationship between main clauses and subordinate clauses must be perceived when the testee chooses appropriate ones. The testee cannot produce any proper sentences without reasoning ability.

Although pictures are excellent devices as stimuli, the testee with poor inferencing ability will have difficulty making a story even if he/she has enough ability to generate sentences. It is difficult to measure a specific ability discretely.

D. Strategies of Language Use

This test is to evaluate a particular language use on a syntactic and stylistic point. This point is one of the strategies of language use. Writing of native "super adults" (K. Hunt referred to voluntary contributors to the Harper's and the Atlantic) is the target for EFL learners. Length and complexity are mutual factors to improve writing. EFL writers should be more conscious of this point.

VII. Reflections on the Process

A. Problems

This test is limited only to two types for convenience. The first intention was to develop the comprehensive test focusing on a particular point, which means to involve many types of test from multiple choice or true or false (to measure receptive knowledge) to creative writing (to measure productive knowledge). Reducing measuring power by symplifying test form must be avoided. Adopting only productive type resulted in receiving a heavy load of scoring.

B. Value

Writing tests utilizing test specifications is indeed helpful to test writers in that objectives, what kind of testees, what skills to be tested and so on are not missed. As a result, tests consistent and congruent to initial objectives are developed.

CONCLUSION

This paper attempted to develop two kinds of test following the particular test specifications. One offers a hint for each test item to facilitate the testee unfamiliar with such tasks to combine sentences. That results in no possible choice. The other offers full freedom in syntactic writing, which enables the test to measure directly the testee's syntactic writing level in the natural environment. These two tests put their place at the extreme opposite positions. Another type of combining test situated in the middle of the other two may be used without any hints after the former test is widely used in order to help L2 learners develop their writing proficiency syntactically.

REFERENCES

Backman, Lyle F. 1990. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.

Brown, J. D. and Hudson, T. 1992c. *Criterion-Referenced Language Testing*. (in press)Cohen, Andrew D. 1980. *Testing Language Ability in the Classroom*. Rowley, Mass.: Newbury House Publishers.

Heaton, J. B. 1975. *Writing English Language Tests*. New York: Longman.

Hudson, T. 1991. Relationships among IRT item discrimination and item fit indices in criterion-referenced language testing, *Language Testing*, 8, 160-181. Edward Arnold.

Hughes, A. 1989. *Testing for Language Teachers*. Cambridge: Cambridge University Press.

Oller, J. W. Jr. 1979. *Language Tests at School*. New York: Longman.

Weir, C. J. 1990. *Communicative Language Testing*.

Test Item Sources

Curtis Kelly and Ian Shortreed (1985). *Significant Scribbles*. Lingual House.

Eiken Ikkyu Niji Shiken Taisaku Kyohon (1990). Nihon Eigo Kyoiku Kyokai.